Broadening access to primary education: contract teacher programs and their impact on education outcomes in Africa – an econometric evaluation for Niger

Bourdon, Jean ; Frölich, Markus ; Michaelowa, Katharina

Abstract: For Sub-Saharan Africa as a whole, but particularly for countries in the Sahel zone, full primary enrolment and completion at acceptable quality as codified in the Millennium Development Goals and the Education for All objectives still remains a major challenge. In order to enhance education supply, many of these countries have launched large scale teacher recruitment programs in recent years, whereby the teachers are no longer engaged in civil servant positions, but on the basis of fixed-term contracts typically implying considerably lower salaries and a sharply reduced duration of professional training. While this policy has led to a boost of primary enrolment, stakeholders in the education system generally fear an important loss in the quality of education. Using data from the "Program on the Analysis of Education Systems" (PASEC) for Niger in 2000/2001, we show that once confounding factors are controlled for, the performance of contract teachers is not generally worse than the performance of other teachers. Matching students taught by contract teachers to those taught by civil servants provides no significant evidence of an advantage of the latter in grade 5. In grade 2, there is evidence for a sizeable advantage of traditional teachers – but only as long as job experience is not appropriately taken into account. Given the strong impact on enrolment and the generally insignificant effect on education quality, the overall assessment of the program remains clearly positive.

Broadening Access to Primary Education: Contract Teacher Programs and their Impact on Education Outcomes in Africa – An Econometric Evaluation for the Republic of Niger –

Paper prepared for the
Annual Conference of the Research Committee Development Economics
of the German Economic Association
on "Pro-Poor Growth"
Kiel Institute for World Economics,
July 8-9, 2005

- revised version, 14 September 05 -

Jean Bourdon, Institute for Research in the Sociology and
Economics of Education (IREDU)
jbourdon@u-bourgogne.fr
Markus Frölich, University College London (UCL)
markus.froelich@unisg.ch
Katharina Michaelowa, Hamburg Institute of International Economics (HWWA)
k-michaelowa@hwwa.de

Abstract
For Sub-Saharan Africa as a whole, but particularly for countries in the Sahel zone, full primary enrolment and completion at acceptable quality as codified in the Millennium Development Goals and the Education for All objectives still remains a major challenge. In order to enhance education supply, many of these countries have launched large scale teacher recruitment programs in recent years, whereby the teachers are no longer engaged in civil servant positions, but on the basis of fixed-term contracts typically implying considerably lower salaries and a sharply reduced duration of professional training. While this policy has led to a boost of primary enrolment, stakeholders in the education system generally fear an important loss in the quality of education. Using data from the "Program on the Analysis of Education Systems" (PASEC) for Niger in 2000/2001, we show that once confounding factors are controlled for, the performance of contract teachers is not generally worse than the performance of other teachers. Matching students taught by contract teachers to those taught by civil servants provides no significant evidence of an advantage of the latter in grade 5. In grade 2, there is evidence for a sizeable advantage of traditional teachers – but only as long as job experience is not appropriately taken into account. Given the strong impact on enrolment and the generally insignificant effect on education quality, the overall assessment of the program remains clearly positive.

JEL classifications: O15, I 21, C14

Acknowledgements
This paper was written while the third author was on a stay at IREDU, on the invitation of the economics department of the University of Burgundy. Support and hospitality of both the institute and the economics department are gratefully acknowledged. Moreover, the authors wish to thank Marie Waller for many helpful comments and an anonymous referee for many complementary suggestions.

1. Introduction

As highlighted in the Millennium Development Goals, enhancing access to primary education is an integral part of poverty reduction. It is as much an objective in itself as a necessary tool for economic development and growth. Governments who fail to provide children and adults in their country with the opportunity to acquire basic knowledge and skills, fail to satisfy basic needs and thus fail with respect to a highly relevant dimension of poverty reduction. As a tool for economic development and growth, the influence of education works both directly, via enhanced productivity, and indirectly, channeled by the influence of education on health, nutrition, reduced population growth and so forth. These linkages are now largely documented in the empirical literature (for an overview, see e.g. Michaelowa 2001a).

While the objective of providing at least a basic education for all children has been reiterated over and over again by the international community, and has been codified not only in the Millennium Development Goals, but also in the Education for All (EFA) objectives in 1990 and in 2000, many countries still face considerable problems in meeting the challenge. These countries are primarily located in sub-Saharan Africa, and the situation is at worst in the francophone countries of the Sahel region. According to Bruns, Mingat and Rakotomalala (2002, Tables 1.1 and 2.4) only 45% of African children complete primary school, and all francophone sub-Saharan African countries except Gabon and Togo show figures that lie below or around this regional average. Out of seven countries with primary completion rates at 25% or below, six are francophone. With primary completion rates below 20% and an overall net primary enrolment rate of only 31%, Niger holds the sad record of the worst performer worldwide at the end of the 1990s (UIS 2004, PASEC 2005a).

In terms of development over time, although there has been notable progress in some countries like Benin, Guinea, Mali and Mauritania, in more than half of the francophone African countries, completion rates have stagnated or even declined over the 1990s. The difficulties faced by francophone sub-Saharan Africa are further exacerbated by current inefficiencies in the allocation of financial resources. In a multivariate regression analysis with regional dummies, Mingat and Suchaut (2000, p. 8) show that the relative disadvantage of francophone, and especially Sahelian countries, even holds when GDP per capita is corrected for.

One major constraint for the supply of public education in these countries has always been the relatively high financial burden of teacher remuneration. In the Sahel countries Burkina Faso, Mali and Niger, towards the end of the 1990s, primary teacher salaries amounted to more than 6 times GDP per capita, as compared to 2-5 times GDP per capita in most other African countries and 1.3 times GDP per capita in OECD countries (MINEDAF 2002, p. 117, OECD 2002, Table D6.1). The most extreme case was Niger, with a ratio as high as 9.6. At such high rates in relation to national income, it was obviously impossible for the country to hire the necessary number of teachers to meet the rising demand for education exacerbated by high population growth. To a somewhat lesser extent, the same applies for the other countries in the Sahel region and a high number of other predominantly francophone African countries.

In order to reverse the trend and to enhance education supply, many of these countries have launched large scale teacher recruitment programs in recent years, whereby the teachers are no longer engaged in traditional civil servant positions, but on the basis of fixed-term contracts typically implying considerably lower salaries and a sharply reduced duration of professional training. Even earlier, parents of school-aged children often resort to private

initiatives, opening their own schools with privately engaged teachers, on contracts at considerably lower rates than those foreseen in the public sector.

Overall, the employment of these "contract teachers" has led to a boost of primary enrolment which has already become evident in the most recent national education statistics. At the same time, stakeholders in the education system generally fear an important loss in education quality. They argue that the relaxation of professional training requirements and the loss in teacher job satisfaction supposedly implied by reduced salaries and less job security will necessarily lead to a collapse of the education system in the long run. Therefore, policy makers have to cope with increasing political pressure, typically organized by teacher unions, if they want to continue to privilege enrolment. Moreover, they are largely uninformed about the impact on educational quality these policies really bring about.

It is the objective of this paper to quantify the effects of the contract teacher statute on education quality in terms of primary student achievement and to find out whether the suggested trade-off between enrolment and education quality really exists. Based on micro data provided by the "Program on the Analysis of Education Systems" (PASEC)[1] from a student and teacher survey carried out in Niger, we will analyze whether the engagement of "contract teachers" truly has a significant impact on student learning. Interpreting the teacher's statute as a "contract teacher" as a "treatment" in the sense of the treatment evaluation literature, we will make use of propensity score matching as suggested by Rosenbaum and Rubin (1983) to compare similar students, differing only with respect to their teachers. On this basis, we will check the robustness of initial results on this question reported in the PASEC country report (PASEC 2005a).

Some earlier studies have attempted to estimate the impact of contract teachers in other countries. PASEC (2003, 2004, 2005b) provide some regression estimates for Guinea, Togo and Mali. Bourdon (2005) reproduces the results for Mali using propensity score matching. Vegas and de Laat (2003) reexamine the PASEC data for Togo. Michaelowa (2001b, 2002) analyzes the evidence from earlier PASEC surveys in Burkina Faso, Cameroon, Côte d'Ivoire, Madagascar and Senegal. These studies will allow us to consider the results for Niger in a broader regional context and to compare national differences of the teacher recruitment programs employed.

This paper is structured as follows: Based on information compiled by Bernard, Tiyab and Vianou (2004) Section 2 demonstrates the quantitative relevance of contract teacher programs and their development over time for both Niger and other countries of the region. Section 3 provides an overview of the data available for the analysis of contract teachers' impact on student achievement in Niger. Section 4 discusses the estimation approach and presents the results. The latter will then be put in perspective and compared to the results of other country studies in order to conclude with some policy recommendations in Section 5.

2. The relevance of contract teacher programs and their development over time

According to data from the Nigerian administration compiled by PASEC (2005a), the country's population more than doubled between 1977 and 2001. Between 1990 and 2003 alone, it increased from 7.7 to 11.5 million. More than half of the population is below 15

[1] Original French title: "Programme d'analyse des systèmes éducatifs de la CONFEMEN", whereby CONFEMEN stands for the Conference of Francophone Education Ministers ("Conférence des ministres de l'éducation des pays ayant le français en partage"). To obtain the original data set, contact pasec@sentoo.sn.

years of age. Thus, population pressure constitutes a considerable challenge for the provision of educational services. Since, at the same time, economic growth has never been able to match population growth, available resources per head have declined, resulting in lower tax revenue per capita and reductions in the degree of freedom for public expenditure. In addition, even as a percentage of GDP, educational expenditure was subject to severe cuts, decreasing from around 4.5% in 1990 to 2.6% in 2002.

The challenge of providing education to a strongly increasing population of children and youth under tight budget constraints is similar for many other countries in the region. According to estimates of UNESCO / Pôle de Dakar, the members of the Economic Community of Western African States (ECOWAS) as a whole faced an increase in primary student numbers of 53% between 1987 and 2001, and will have to cope with a further increase of 83% between 2001 and 2015 in order to meet the EFA objectives. If class size is given, this results in an equally important need for additional teachers (Amelewonou, Brossard and Gacougnolle 2004). Theoretically, an alternative could be a further increase of the student-teacher ratio, but at a given average class size of over 45 (with often much higher class sizes, especially in urban areas) this option is generally considered as unviable.

However, given tight budget constraints, it became obvious towards the late 1990s that engaging more teachers on traditional civil servants contracts with salaries and other benefits was illusionary as well. These salaries were far above the market rates as they originated in the French law on the equality of remuneration for civil servants voted on the initiative of Lamine Gueye (native from Senegal) in 1950.

In order to be sustainable, an education system must be able to exist on the country's own resources, at least in the long run, i.e. expenditure must remain in a viable proportion to national income. From this perspective, traditional civil servant teacher salaries in many Western African countries, notably in the Sahel region, appeared to be clearly unsustainable.

With primary teacher salaries amounting to about 9 times GDP per capita in the late 1990s, teacher remuneration in Niger as compared to per capita income was the highest worldwide (see Table 1, column 1). In 1998, the country therefore decided to introduce a new teacher recruitment system whereby new teachers would no longer be integrated in the civil service, but employed on fixed term contracts with a salary about half as high as traditional teacher salaries.[2] While other countries followed similar policies, the policy shift in Niger was particularly marked and brought about the massive employment of new teachers. After 1998, almost 2800 new teachers were recruited every year, as compared to an average of 520 per year between 1990 and 1998 (PASEC 2005a), so that in the year 2000, public "contract teachers" already made up 50% of the entire primary teacher population (see right hand side of Table 1). In other countries, public authorities were generally more reluctant, but private stakeholders such as parents and local communities often took over the initiative.

[2] Obviously other ways of reducing the financial burden of teachers salaries would also have been conceivable, in particular a reduction of civil servants' remuneration could have been considered. However, it is generally perceived as politically very difficult or even impossible to change existing rights and regulations for civil servant already in office (insider-outsider problem).

Table 1: Distribution of primary teachers according to their statute

Country	Teacher remuneration			Distribution across statutes		
	Civil Servants	Contract teachers		Civil servants	Contract teachers	
		Public	*Private*		*Public*	*Private*
Benin (2002)	5,2	2,1	1,3	55%	16%	29%
Burkina Faso (2002)	5,8	5,6	2,2	64%	24%	12%
Cameroon (2002)	5,3	1,4	0,8	35%	20%	45%
Chad (2002)	8,2	-	2,3	32%	0%	68%
Congo (2003)	2,4	0,9	0,6	42%	4%	54%
Côte d'Ivoire (2001)	4,8	-	-	87%	0%	13%
Guinea (2000)	3,5	1,1	-	52%	30%	18%
Mali (2000)	5,8	1,5	0,9	71%	8%	21%
Niger (2000)	8,9	3,5	-	46%	50%	4%
Senegal (2003)	5,7	2,6	-	44%	41%	15%
Togo (2001)	6,4	3,3	1,3	35%	30%	35%
Mean	**5,6**	**2,4**	**1,3**	**51%**	**20%**	**29%**

Notes: *Public*: under contract with public authorities; *Private*: under contract with parents or local communities. On average, in OECD countries, primary school teachers' salary corresponds to 1.3 times GDP per capita; in Germany and UK the factor is 1.5, and in the US, it is 1.1 to name but a few examples. (However, one should be cautious with direct comparisons given the scarcity of human capital in developing countries.)

Source: World Bank, Africa Region (quoted from Bernard, Tiyab and Vianou 2004, p. 5), OECD (2002, Table D6.1).

While the purpose of introducing the statute of "contract teachers" has always been to further enrolment at reduced cost, the concrete implementation of this policy has been quite different in each country. While some countries like Niger and Togo started to employ all new teachers on fixed term contracts, countries like Mali continued to engage a limited number of civil servants. Moreover, there are considerable cross-country differences in the relative salary of the new contract teachers. As depicted on the left hand side of Table 1, in some countries like Cameroon and Mali, public contract teachers receive only about 25% of the traditional teacher salaries. In other countries the discrepancy is much less extreme, in particular in Burkina Faso where one can hardly make out a real difference. Privately employed contract teachers earn much lower salaries than traditional teachers in all countries, and often even much less than the new public contract teachers do.

Further differences exist with respect to entry requirements in terms of educational attainment and professional training. Typically, professional training has been considerably reduced from several years in specialized teacher training institutes ("Ecoles Normales") to a few months, or even weeks provided by diverse institutions or as on-the-job training under the mentorship of senior teachers. This reduction in teacher training also reduced the cost incurred by the education system, in particular as teacher candidates often received scholarships during this training period, i.e. before being actually employed in schools.

In some countries like Guinea, cuts in the duration of professional training have gone hand in hand with a redefinition of course content and an attempt to adjust the curricula to issues truly relevant to teaching practice. A mandatory training period of 15-18 months including 6-9 months of practical teaching experience under the supervision of a senior colleague was built

into the new recruitment procedure (Faoura 2004, pp. 105ff.)[3] In Mali, teacher candidates who have not attended other forms of pedagogical training, have to follow a three-month course preparing them for their work (PASEC 2005b). As opposed to the situation in Guinea and Mali, contract teachers in Niger only receive a 45-day training, and even the latter requirement is not always met in practice. While some contract teachers have effectively attended longer training courses, they can be assumed to have been candidates for traditional teaching positions who have not been able to find a position on the labor market. Traditional civil servant teachers had to follow a training administered in the "Ecole Normale" for one or two years.

Similar cross-country differences exist with respect to educational attainment required to enter the teaching profession. While in Guinea and Mali, a higher secondary completion exam ("baccalauréat") is required as a minimum, contract teachers in Niger may also start from a level of completed 10^{th} grade ("BEPC") if they clear the entrance exam.

All of these factors seem to put Niger in a particularly difficult situation with respect to potential impacts of its contract teacher program on education quality. These impacts will be considered in more detail in the following sections. Theoretically, we will have to consider the following potential effects: (1) an incentive effect of the teaching contract, (2) a selection effect (changed demand for and supply of new teachers), and (3) a dynamic effect.

The direction of the incentive effect cannot be determined theoretically: On the one hand, the unfavorable conditions of new teacher contracts could be regarded as unfair and demotivating, and short-term contracts could prevent personal investments in pedagogical training and school specific human capital. On the other hand, for contract teachers, further employment prospects depend on performance and, among other things, parents' satisfaction, so that from this perspective, the contract statute could be expected to have a positive incentive effect.

As far as the selection effect is concerned, the changed employment conditions could lead to a different composition of teacher candidates. On the one hand, we would expect a lower number of highly skilled candidates due to the inferior contract conditions. On the other hand, the reduced entry requirements could reduce entry costs and increase the attractiveness of (temporary) teaching positions. The higher demand for teachers would lead us to expect a lower quality of the marginal (newly employed) teacher.

The dynamic effect, finally, refers to a potential change of teacher behavior over time. In particular, the inferior contract conditions may induce a reduced retention period of teaching staff. This effect could lead to a different distribution of job experience before and after the reform.

As we will see below, the data at hand will allow us to estimate a combination of the overall incentive and selection effect without, however, allowing us to distinguish between these effects or between any of the underlying indirect effects. At the same time, a consideration of the dynamic effect cannot be provided so far because the distribution of job experience in the teaching population has not yet reached its equilibrium due to the relatively recent introduction of the new program, and is thus truncated from above.

[3] The revision of the training program was actually at the center of reforms in Guinea. The new contract teachers are known under the name of the training program "Formation initiale des maîtres de Guinée" (FIMG) – a pilot program initiated by the World Bank - rather than as "contract teachers" (PASEC 2003).

3. The PASEC data set for Niger and selected characteristics of contract teachers, their students and schools

The evaluation of the impact of the contract teacher program in Niger can be based on a stratified survey of schools and students in the 2nd and 5th grade of primary education during the academic year 2000/2001. The impact of the program will thus be measured on the basis of student achievement. Sampling was carried out in a way that each inspection, i.e. each school district, should be covered, and within these districts, half of the schools were drawn randomly from those schools which comprise the 2nd and the 5th grade level. If several classes existed within a grade level of a given school, one class was randomly selected. Depending on whether the teacher of the classes drawn in this way was a civil servant or a contract teacher, a nearby school with similar characteristics was selected to provide evidence for the other type of teacher statute. As, however, it was not always possible to find the opposite category for both grades simultaneously in the same school, the final sample is equally spread between contract teachers and teachers on traditional civil servant positions only for grade 2. In grade 5, we find twice as many civil servants as contract teachers. To a certain extent, this may reflect the real proportions, as contract teachers tend to be employed more often for lower than for higher grades. All in all, the sample includes 280 classes (of which 140 in 2nd and 140 in 5th grade) within 26 inspections. In each class, 15 students were randomly selected to take a test in French and Mathematics. These tests were designed by PASEC for assessing educational standards in francophone Africa and are not used by the Nigerian authorities for any official purposes, e.g. teacher assessments. These standardized tests are oriented at the typical curriculum of francophone African primary education. Both tests were administered in French language. The Math test contains a wide variety of items ranging from calculus over problem solving (application to situations of daily life) to simple geometry. The French test covers general understanding and orthography as well as grammar skills. Tests were administered in the classroom, item by item, following detailed instructions on the way to present each question and the time to be allocated to its response. Students were tested at the beginning and at the end of the academic year. Test results for both pre- and post-test are coded in terms of the percentage of test items answered correctly in each of the two subjects French and Math. The average test scores in grade 2 are 42% for French and 41% for Math, with a standard deviation of 25% and 26%. In grade 5 average test scores are 28% for French and 31% for Math with a standard deviation of 15% and 16% respectively.[4]

In addition, students were interviewed on their personal characteristics (gender, age), their family background (e.g. family possessions, parents' literacy, use of the French language at home) and their prior educational background (e.g. PASEC scores at the beginning of the academic year, grade repetition). Moreover, rich data on teachers and schools was collected using separate teacher and director questionnaires. Finally, this data was complemented by information provided by Niger's ministry of basic education and alphabetization (MEB_1A 2004) on diverse regional characteristics and characteristics of the school districts (such as the regional enrolment rate, the growth of the regional school aged population etc.).

Table 2 provides an overview of some selected characteristics of students, teachers, classrooms, schools and regions. The columns show the overall mean of the respective variable, as well as the averages for classes taught by contract teachers and civil servant teachers respectively. The total number of observations is less than 2100 (=140×15) for each grade level because in some (rare) cases, teacher and director questionnaires were not duly filled in. In other cases, the students' final scores are missing due to their absence on the day

[4] These values differ from those presented in Annex 3 because in this section, the values refer to the distribution of students while in the annex, they refer to the distribution of classes.

of the posttest. These missing values at the student level account for the bulk of missing observations. As Niger is an extremely poor country, the high rate of absence on the day of the final exam can probably be explained to a large extent by early drop-out, by sickness, or by household and harvesting activities on the day of the evaluation. Simple correlation indicates that the occurrence of these missing values is independent of the teacher's statute as a civil servant or a contract teacher.

Generally, when full questionnaires or test results were missing, the observations concerned had to be deleted from the data set. Whenever only information for individual variables was missing, however, missing values could generally be imputed using related questions taken from the same or other questionnaires. If information from several other variables could be included for imputation, imputation was carried out using linear regression on all of these variables. In addition, missing value indicators were generated and included in subsequent estimations. In principle, including missing value indicators in the empirical analyses renders the imputation irrelevant. However, several of these missing value indicators are highly, but not perfectly, correlated, leading to collinearity problems in later regressions. Therefore, not all missing value indicators can be included, such that plausibly imputed values become important. The imputed values also permit to reduce the covariate space in that most of the missing value indicators turn out to be highly insignificant in subsequent regressions. This instills confidence in the fact that students with missing items do not represent a very selected subpopulation.

A comparison of test scores in Table 2 indicates that on average, in the sample for both 2nd and 5th grade, students taught by teachers on civil servant contracts obtained higher test scores. However, these differences are only partly significant (at the 5% level), and furthermore, they may be related to the factors other than the teachers' contract statute. They may, for instance, be a consequence of the assignment of contract teachers to different learning environments, or to different characteristics of the teachers themselves.

The general comparison of mean values for various context variables suggests that bias in the perceived difference between contract and civil servant teachers may partly arise from the fact that the former seem to be employed primarily in classes or schools where students already had lower scores at the beginning of the academic year. Table 2 shows that students' initial scores in French and Math differ about as much between the two groups as the final scores measured at the end of the year. Thus, looking at progress rather than the final outcome alone, the performance of the traditional teachers does not seem to be much different from the performance of contract teachers.

Moreover, there are important differences in teacher characteristics which should have been expected given the relatively recent introduction of the program and the legal constraints related to its implementation. In particular, as only adults between 18 and 35 years of age are eligible for employment as contract teachers, at the time of the survey, the maximum age of a contract teacher could be 39. For this reason, on average, contract teachers in the sample are much younger than their peers in the civil service. The same situation is reflected in teachers' job experience. While a few teachers seem to have been employed on related tasks for some years before they became contract teachers, the maximum job experience for this group in the sample is 8 years. Among traditional teachers, those with job experience of less than 4 years are rare exceptions. Theoretically, no civil servant teachers should have been engaged after the introduction of the contract teacher program at all, but in some rare cases, this rule seems to have been violated, or alternatively, some teachers may have interrupted their work for

Table 2: Selected teacher, school and student characteristics, by teachers' statute and grade, 2000/2001

	Grade 2			Grade 5		
Variables	**All teachers**	**Civil servants**	**Contract teachers**	**All teachers**	**Civil servants**	**Contract teachers**
Number of students	1881	940	941	1905	1494	411
Number of classes	123	60	63	121	93	28
Student characteristics						
Test scores (in % of correct answers, averages)						
Final score French	42.3 (1.65)	43.7 (2.50)	40.0 (2.15)	28.5 (0.88)	29.2 (1.05)	25.9 (1.40)
Final score Math	40.6 (1.48)	42.3 (2.41)	38.4 (1.74)	31.3 (0.90)	32.2 (1.05)	28.1 (1.57)
Initial score French	16.5 (1.23)	16.8 (1.66)	15.2 (1.80)	25.4 (0.86)	26.0 (0.97)	23.3 (1.77)
Initial score Math	34.6 (1.92)	36.2 (2.57)	31.5 (2.78)	29.1 (0.92)	29.9 (1.05)	26.5 (1.85)
Other student characteristics						
Index of household possessions (1-8)	0.53 (0.04)	0.55 (0.06)	0.52 (0.06)	0.56 (0.04)	0.52 (0.05)	0.68 (0.11)
Index of household tasks in which student is involved	3.50 (0.12)	3.46 (0.18)	3.57 (0.17)	4.30 (0.12)	4.23 (0.14)	4.55 (0.22)
Share of students with books at home	n.a.	n.a.	n.a.	0.28 (0.02)	0.28 (0.02)	0.27 (0.04)
Share of students with French book	0.20 (0.02)	0.18 (0.02)	0.22 (0.02)	0.28 (0.02)	0.28 (0.02)	0.26 (0.03)
Share of students with Math book	0.13 (0.01)	0.13 (0.02)	0.13 (0.02)	0.16 (0.01)	0.17 (0.02)	0.16 (0.02)
Share of literate mothers	n.a.	n.a.	n.a.	0.32 (0.02)	0.29 (0.02)	0.40 (0.04)
Share of literate fathers	n.a.	n.a.	n.a.	0.55 (0.02)	0.54 (0.02)	0.58 (0.04)
Share of students speaking French at home	0.06 (0.01)	0.06 (0.02)	0.06 (0.02)	0.09 (0.02)	0.09 (0.02)	0.09 (0.04)
Share of students speaking Arab at home	0.02 (0.01)	0.02 (0.01)	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)	0.01 (0.00)
Share of students speaking Tamacheq at home	0.02 (0.00)	0.02 (0.01)	0.02 (0.01)	0.02 (0.00)	0.02 (0.00)	0.05 (0.01)
Teacher characteristics						
Age (in years)	30.6 (0.53)	33.7 (0.75)	27.6 (0.52)	33.0 (0.53)	34.2 (0.58)	28.9 (0.86)
Job experience (in years)	6.26 (0.61)	10.27 (1.01)	2.44 (0.16)	9.45 (0.71)	11.4 (0.81)	2.82 (0.34)
Index of teachers' educational attainment (0-6)	3.78 (0.08)	3.82 (0.12)	3.75 (0.12)	4.09 (0.09)	4.05 (0.10)	4.21 (0.18)
Duration of teacher training index (0-4)	1.73 (0.10)	1.62 (0.13)	1.84 (0.15)	1.61 (0.08)	1.72 (0.10)	1.25 (0.10)
Classroom and school characteristics						
Class size	44.2 (1.33)	44.8 (2.15)	43.6 (1.61)	45.1 (1.33)	44.1 (1.54)	48.5 (2.58)
Share of schools with active parents	0.69 (0.04)	0.63 (0.06)	0.75 (0.06)	0.72 (0.04)	0.76 (0.04)	0.57 (0.10)
Share of schools participating in pilot projects	0.37 (0.04)	0.33 (0.06)	0.40 (0.06)	0.36 (0.04)	0.42 (0.05)	0.18 (0.07)
Index of school equipment (0-10)	3.84 (0.21)	3.85 (0.29)	3.83 (0.31)	3.99 (0.22)	3.87 (0.26)	4.39 (0.41)
Share of classes where blackboard and chalk available	0.91 (0.03)	0.88 (0.04)	0.94 (0.03)	0.93 (0.02)	0.90 (0.03)	1.00 0.00
Index of directors' educational attainment (0-6)	3.99 (0.10)	3.90 (0.14)	4.08 (0.15)	4.01 (0.10)	4.05 (0.12)	3.86 (0.22)
School location in a rural area	0.34 (0.04)	0.30 (0.06)	0.38 (0.06)	0.32 (0.04)	0.30 (0.05)	0.39 (0.09)
Regional characteristics						
Population growth rate, boys 7-12	0.62 (0.02)	0.61 (0.03)	0.62 (0.02)	0.59 (0.02)	0.57 (0.02)	0.65 (0.04)
Share of buildings in bad condition	38.6 (0.81)	39.3 (1.15)	38.0 (1.14)	37.9 (0.87)	38.3 (1.01)	36.5 (1.72)
Share of multi-grade schools	0.08 (0.01)	0.08 (0.01)	0.09 (0.01)	0.08 (0.01)	0.09 (0.01)	0.06 (0.01)

Note: Standard errors in parenthesis. At student level, clustering into classes is taken into account.

some years after having been initially employed before 1998. All in all, there remains a sharp difference of almost 8 years between the average job experience of contract teachers and civil servants. If job experience has a positive impact on student achievement, this difference may lead to an important underestimation of contract teachers' performance.

With respect to teachers' educational attainment and the duration of their pedagogical training, differences are less pronounced than what might have been expected. Despite the fact that no baccalauréat is required for entry, in 5th grade, contract teachers actually tend to be more highly educated than traditional civil servants. This may be due to the saturation of the labor market and a certain oversupply of upper secondary graduates who are willing to reorient their professional objectives towards teaching, at least until other job opportunities open up for them. Overall, the differences appear to be insignificant.

As far as pedagogical training is concerned, for 5th grade teachers, we do find the expected significant differences. Only one of the contract teachers has gone through more than the minimum training required. At 2nd grade, however, several contract teachers recruited on the new program seem to have followed the traditional teacher training some time earlier. At the same time, both in 2nd and in 5th grade, the majority of traditional teachers do not meet the official employment criterion of at least one year of pedagogical training. The value of the training index utilized in Table 2 is above 3 if a teacher effectively spent at least one year in a training course, while the average for civil servant teachers does not even reach 1.8 in practice. In many cases, this regulation must have been waived so that the differences between contract teachers and civil servants are less pronounced than expected.

With respect to other descriptors of students' learning environment, only a slight difference can be made out between the two groups of teachers, especially in 2nd grade. This probably reflects the effort to survey a nearby, "similar" school with a civil servant teacher for each school surveyed with a contract teacher. By doing so, the sampling procedure attempted to replicate a truly experimental situation as closely as possible. Nevertheless, for some variables, some differences remain, notably in 5th grade where – due to the constraint to survey classes from the same school as in 2nd grade – the appropriate pair of teachers could not always be selected. In this context, we note in particular, that 5th grade contract teachers tend to teach in regions with higher pressure of population growth and in classes with more students, but that students' family environment tends to be more favorable (higher index of household possessions, higher share of literate mothers). In the following, we will examine whether, once these confounding factors are taken into account, the apparent superiority of traditional teachers can still be supported.

4. Evaluating the effect of the contract teacher program on the basis of propensity score matching

In order to control for confounding factors and to identify the true effect of the contract teacher program, we try to match classes taught by contract teachers with classes taught by civil servants. Following the terminology of the treatment evaluation literature, the type of teacher (contract teacher or civil servant) is considered as a "treatment" to which the class is subjected. The empirical strategy attempts to select control classes corresponding in their characteristics to those classes who received the treatment. If all variables simultaneously influencing the treatment decision and student achievement are taken into account, the "selection on observables" (Heckman and Robb 1985) or "conditional independence" (Lechner 1999) assumption is satisfied and the impact of the contract teacher statute can be

identified. Let Y^0 denote the average test score outcome in a class taught by a regular teacher and let Y^1 denote the test scores outcome if this class were taught by a contract teacher. The difference Y^1-Y^0 in these potential outcomes represents the gain or loss in achievement if a regular teacher were to be replaced by a contract teacher. Let X denote all the confounding variables which simultaneously influence these potential outcomes and the probability that the teacher is employed under this non civil-servant contract, including student, teacher, classroom, school and regional characteristics. As our data set is very rich and includes all variables typically assumed to be relevant for student achievement, we are confident that the CIA is satisfied. By conditioning on X, the potential outcomes are then identified as:

$$E[Y^1|X] = E(Y|D=1, X) \quad \text{and} \quad E[Y^0|X] = E(Y|D=0, X),$$

where D denotes the observed contract status of the teacher (D=1 for contract teachers, D=0 for regular teachers). With this relationship and by averaging with respect to the population distribution of X, the average treatment effect on the treated[5] is identified as

$$(1) \qquad E[Y^1-Y^0] = \int [E(Y^1|X) - E(Y^0|X)] \cdot dF_{X|D=1} = \int [E(Y|D=1,X) - E(Y|D=0,X)] \cdot dF_{X|D=1} .$$

This matching approach is preferred to traditional regression analysis for two reasons: First, in addition to the conditional independence assumption, regression analysis requires the assumption of particular functional relationships. Generally, for convenience, the functional form is assumed to be linear which actually seems rather implausible when inputs into the educational production process are concerned. Second, on the basis of this hypothetical functional relationship, regression analysis expands its predictions into areas for which no appropriate control observations are available. It thereby tends to compare classes which, under closer scrutiny, one would be very reluctant to compare. In regression analysis, the problem of comparing the incomparable does not become apparent, however, as it is simply assumed away. Given the differences between the treated and the control group highlighted in the previous section, this could lead to a considerable distortion of results. Therefore, ensuring that characteristics of those observations compared with each other are in a common range (i.e. have common support) should be an explicit matter of concern.[6]

As demonstrated by Rosenbaum and Rubin (1983), the estimation of the treatment effect can be facilitated if the information incorporated in the relevant control variables is first projected into a single variable, the propensity score p(x)=P(D=1|X=x). They showed that if matching on X is consistent, matching with respect to the propensity score p(x) is consistent as well. The multidimensional problem of matching on X is thereby reduced to the one-dimensional problem of matching on p(X). The propensity score can be estimated e.g. by a probit regression of the binary treatment variable "contract teacher status" on X. In the context of this preliminary estimation of a binary treatment variable, results are typically more robust with respect to different distributional assumptions than they are at the level of the final (continuous) outcome variable Y (see e.g. Black and Smith, 2004, footnote 13). Moreover, given that at class level our sample only includes a relatively small number of observations, a parametric approach for the estimation of the propensity score is useful to improve the

[5] See Section 4.1 for an explanation why we focus on the average treatment effect on the treated (ATT) rather than on the average treatment effect on the untreated (ATU) or the average treatment effect (ATE).

[6] A third advantage is that the matching approach relaxes the assumptions about the independence of the error term. Whereas in a traditional linear model approach it must be assumed that the error term is independent of the regressors, i.e. E[u|X,D]=0, this assumption may be relaxed here to E[u|X,D] = E[u|X], i.e. the error term could still be correlated with the regressors X.

robustness of results with respect to other variations in model specification. We therefore follow the general trend in the literature to use a probit model for the estimation of the propensity score.

For the estimation of the conditional expectation function we avoid parametric assumptions and adopt Nadaraya-Watson kernel regression using a Gaussian kernel with a bandwidth of 0.06 as suggested by Leuven and Sianesi (2003). While nearest neighbor matching or radius matching might come conceptually closer to the original idea of comparing only classes with identical (or very similar) characteristics, kernel matching makes more efficient use of the available information as it does not discard any observations, but weighs the observations according to the appropriateness of the match.

4.1. Ensuring common support

As mentioned above, given that teachers under the new contract teacher program are necessarily different from teachers in civil service with respect to a number of criteria, ensuring common support requires particular attention. As discussed in Section 3, finding similar characteristics in both groups is less problematic than expected for the education and training variables, but highly problematic for teachers' job experience and age. As mentioned earlier, the maximum age and job experience among contract teachers is 39 and 8 years, respectively, while civil servant teachers are generally older and more experienced.

We will pursue two different approaches to deal with this problem. In the first approach (specification 1), neither teacher's job experience nor age is controlled for. This is justified if experience and age are no relevant determinants of students' educational achievement – an assumption for which we find some supporting evidence. Alternatively, this estimate can be interpreted as the treatment effect for a student when replacing a civil servant teacher with one of the new contract teachers. Since the new contract teachers, by the implementation of the reform, must have on average lower age and experience, this measures the *combined* effect of changing contract status and lowering experience. If the effect of experience on educational achievement was indeed zero, this would correspond to the effect of a change in contract status only.

In the second approach (specification 2) to the support problem, we restrict our analysis to the population of teachers with at most 8 years of job experience. In other words, all civil servant teachers with more than 8 years of job experience (and a correspondingly high age) are deleted, which leads to a considerable reduction in sample size. While age can be discarded as insignificant, job experience is then controlled for in the matching estimator and the estimated effect can be interpreted as replacing a civil servant teacher with a contract teacher with the *same* job experience. Hence, this effect differs from specification 1, where the combined effect of changing contract status and lowering experience was estimated. However, these two estimates differ further in that they refer to different populations, which might yield different estimates even if job experience had no effect on educational achievement.

For the first approach, in an initial step, we examine whether teacher job experience and age actually affect student achievement. If it could be established that they are only related to the teacher's statute but not to achievement, omitting them from the model would not induce a bias. In the literature, teachers' age is generally not considered as a relevant determinant of student achievement, but teacher job experience is. Typically, it appears to be relevant in a non-linear way, with a positive effect in the initial years which diminishes over time (see e.g. Michaelowa 2001b, Bernard 1999, and Rivkin, Hanushek and Kain 2005). However, a closer

look at the data for Niger suggests that the situation here might not necessarily be the same. In a simple OLS regression, we find a positively significant impact of job experience only for 5th grade achievement, and this effect appears to be driven exclusively by some strong performers among teachers with extensive job experience. Annex 1 shows a partial scatter plot for student achievement and teacher job experience for the subgroup of voluntary teachers. It suggests that in 5th grade, only three schools (labeled 91, 120 and 121) with teachers with more than 20 years of job experience are responsible for the overall positive relationship. Dropping these outliers from the regression, no impact of job experience can be observed any more. In 2nd grade, the relationship is not significant in the first place.

Based on this preliminary analysis we may conclude that there is no evidence for job experience being a truly relevant predictor of student achievement in the context of primary schools in Niger, so that once the three outliers are eliminated from the sample, we do not need to consider this variable any further. This will be the assumption for our specification 1.

In the second approach (specification 2), we include job experience as well as the other variables which are systemically related to the introduction of the contract teacher statute (variables on teacher's education and training). As mentioned above, this requires the deletion of all observations for teachers with job experience > 8 years. If the additional control variables are really irrelevant, their introduction should not considerably alter the overall results. However, due to the restriction of the population, results will only refer to the subpopulation of relatively young (i.e. inexperienced) teachers.

In addition to job experience, one more characteristic in our model strongly influences student achievement and, in 5th grade, is not represented among contract teachers at all. It is a matter of class management, the so called "double shift" teaching. Double shift teaching describes a situation where two classes use the same classroom at different times of the day, whereby the teacher may be the same or a different one for the two groups. This can be observed most frequently in densely populated urban areas lacking adequate infrastructure (and a sufficiently high number of teachers), and often leads to less hours of effective teaching and reduced learning outcomes (Michaelowa 2001b). Closer scrutiny of the data reveals that in Niger, double shift teaching appears to be extremely rare in the higher grades of primary education, so that even for civil servant teachers, we only find one observation in grade 5. For this reason, the most obvious solution to the support problem is to simply delete this one observation from the 5th grade data set. In 2nd grade, double shift teaching is more widely spread and can be included in the control variables without any problem.

The preliminary adjustments of the data set allow us to eliminate the most obvious sources of the common support problem. Nevertheless, it remains that for certain combinations of these control variables it may be impossible to find appropriate matches. This can be taken care of later on by simply imposing a common support restriction on the propensity score as a whole. In this context, it should be noted that the observations we may have to delete from the data set depend on whether we are merely interested in the treatment effect on the treated, the treatment effect on the untreated, or the average treatment effect reflecting both. Restricting our analysis to the average effect of the treatment on the treated (ATT), we have to be cautious about the inclusion of observations on contract teacher classes whose characteristics lie outside the range of observable characteristics for the control group. However, if we were interested in the effect a contract teacher would have on those classes who have effectively been taught by a civil servant (average treatment effect on the untreated, ATU), we would need to ensure that characteristics of civil servant teacher classes stay within the range of characteristics observed in contract teacher classes. This condition appears to be much more

difficult to meet since our data show more variation in the characteristics of civil servant teachers than in the characteristics of contract teachers. We will therefore restrict our discussion to the average treatment effect on the treated. From a policy point of view, this also appears to be most interesting since the relevant policy question is not whether active civil servant teachers should be replaced by contract teachers, but whether the newly engaged contract teachers show an acceptable performance or whether their statute needs to be changed back into the one of a civil servant.

4.2. Determining relevant control variables and estimation of the propensity score

To determine the relevant control variables, the usual approach is to carry out repeated regressions of the binary treatment variable on various sets of regressors and to keep the most significant variables in the final specification. As our treatment variable is defined at the class level, the number of observations is relatively limited and starting regressions with a very large number of regressors becomes problematic. In particular, several combinations of variables for which observations are very scarce, lead to perfect predictions of the probit model. We therefore adopt a different approach starting with a closer look at the potential determinants of student test scores.

Estimation of education production functions has become a focus of many studies in recent years including numerous studies for developing countries ever since micro survey data have become available in the late 1990s. A recent summary is provided in the EFA Global Monitoring Report 2005 on education quality (UNESCO 2004). The variables presented in Section 3 (Table 2) represent a collection of plausible predictors drawn from this literature. We take these variables (or some composite indicators based on the information of these variables) as the basis for the selection of the relevant variables for our model for Niger. The final selection is then determined through an iterative process between provisional OLS regressions of student achievement and probit regressions of the contract teacher statute. Like the probit regressions, regressions of student achievement are carried out merely at the class level, and class averages are computed for variables originally measured at the level of the individual student. As a general rule, variables are retained for the final estimation of the propensity score if their p-values are below 20% in both the achievement regression and the regression of the teacher statute.[7] We always retain the pre-test scores, even if they are insignificant in the teacher status regression, as they are very important determinants of the final achievement outcomes and incorporate the value-added approach to measuring the impact of teacher statute.

As discussed in Section 4.1, in the second specification, the sample is reduced to classes of teachers with no more than 8 years of job experience. This allows us to introduce job experience as an additional control variable without facing insurmountable problems to ensure common support. In this second specification, we also include teachers' educational attainment and the duration of teacher training. As these variables are structurally linked to the contract teacher statute – although to a lesser extent than we had expected according to formal regulation – and as they are generally believed to be (and frequently empirically established as) relevant predictors of student achievement, it may be more prudent to keep

[7] Starting with the achievement regressions, a stepwise backward selection procedure was used with a significance level of 20% for removal from the model. When a forward selection procedure was carried out alternatively, no additional variables were included in the model and the total number of regressors was reduced. To be sure not to omit any relevant determinant of student achievement, we used the more comprehensive model resulting from the backward selection.

these variables and not rely on simple OLS achievement regressions as a reason for eliminating them from the model.

By reducing the sample and adding the additional explanatory variables, some predictors used in specification 1 do not appear to exert any relevant influence any more. As their p-values are far beyond the 20% cut-off point, they are no longer included in specification 2. Alternative regressions retaining these variables in the model yield very similar results.

It should be noted that we face some problems specifying the model for grade 2 because some potentially relevant explanatory variables (parents' literacy and the availability of books at home) are not available for younger students. This could lead to biased estimates of the treatment effect. Assuming that the average socio-economic background of students in the same school does not vary substantially across grades,[8] we thus use the available information for 5th grade students in the same schools as a proxy for the unobserved 2nd grade variables. One of these proxies is retained as a relevant predictor in the final specification.

Table 3 presents the final selection of variables and their impact on the probability to be taught by a contract teacher, i.e. the results of the probit regressions used to construct the propensity score. It can be observed that unless teacher job experience and training are included in the model, the overall explanatory power of these regressions is very low, especially in grade 2 (pseudo R^2=0.06). In fact, in specification 1 for grade 2, only parents' literacy and the average number of textbooks per student turn out to be significant at the level of 10%. It appears that 2nd grade students in the contract teacher classes included in our sample often come from relatively disadvantaged family backgrounds (in terms of literacy) while at the same time, they are relatively well equipped in terms of textbooks. In 5th grade, the pseudo R^2 is still not very high (18%), but parents' literacy, the share of Tamacheq speaking students, the participation of the school in a pilot project, the location of the school in a rural area and the number of multi-grade schools in the region, all seem to play a certain role in determining the statute of the teacher by whom students are taught in a particular school. To understand the role of Tamacheq it should be noted that this is the language spoken by the nomadic population in the northern arid areas of the country who generally also show a strong learning disadvantage, particularly in grade 5.

[8] More precisely, we only need to assume that *differences* between schools with respect to these variables are unaffected by grade. This remains a plausible assumption even if we consider that at higher grades, the proportion of well-off students may increase due to drop-out being positively correlated with poverty.

Table 3: Estimation of the propensity score[1,2]

Probit estimates	**Grade 2** Specification 1		Specification2		**Grade 5** Specification 1		Specification2	
Dependent variable: Teacher is a contract teacher	Coef.	P>|z|	Coef.	P>|z|	Coef.	P>|z|	Coef.	P>|z|
Initial score, average Math and French	-0.01	0.15	-0.02	0.15	-0.02	0.22	-0.02	0.37
Index of parents' literacy (0-2)[3]	**-0.68**	**0.06**	**-1.32**	**0.01**	**0.96**	**0.04**		
Average number of textbooks per student (0-2)	**0.95**	**0.04**	**1.82**	**0.01**				
Share of Tamacheq speaking students					**5.78**	**0.01**	**9.92**	**0.07**
School is participating in a pilot project					**-0.66**	**0.04**	**-0.81**	**0.06**
Class size	0.04	0.20						
Class size squared	-0.00	0.12						
School is situated in a rural area					**0.85**	**0.03**		
Share of multi-grade classes in the region					**-6.35**	**0.06**	-6.68	0.18
Teacher's job experience (in years)			**-0.64**	**0.00**			**-0.32**	**0.01**
Index of teacher's educational attainment (0-6)			-0.28	0.13			-0.36	0.16
Duration of teacher training index (0-4)			**0.38**	**0.05**			**-0.54**	**0.04**
Constant	-0.08	0.92	**3.83**	**0.00**	-0.76	0.23	**4.43**	**0.01**
N	123		97		117		69	
Pseudo R^2	0.06		0.38		0.18		0.37	
LR, Prob > chi^2	0.06		0.00		0.00		0.00	

[1] Bold figures indicate significance at the 10% level.
[2] In each grade, the same specifications were used to calculate the propensity score for Math and French separately. In this case, the joint pre-test score was replaced by the initial score in the respective subject matter. Results are presented in Annex 2, Tables A2.1 and A2.2.
[3] For grade 2 this variable does not exist. However, assuming that parents' literacy level does not differ much within a given school, the missing information was approximated by information available for grade 5 (see also footnote 5).

Given the information about the survey design provided in Section 3 and the discussion of Table 2 thereafter, the generally low explanatory power of specification 1 in the probit regressions should not come as a surprise. It basically confirms that right from the beginning, the sampling process ensured data collection in a way that came close to data collection from an experimental design. As in 2nd grade for each class with a contract teacher another "similar" class with a civil servant teacher was selected from a nearby school, our probit results confirm that the underlying initial assessment of "similarity" was based on the relevant criteria so that very few of these are significant in ex post regression analysis. For grade 5, the situation is somewhat different because at the sampling stage, appropriate matches could not be found to the same extent as in grade 2.

However, only if teacher job experience, educational attainment and training are taken into account (specification 2), does the situation change considerably. As these variables are directly related to the teacher statute rather than to the school, they could not be considered when searching for a "similar" school at the sampling stage. While educational attainment appears to be less relevant, teacher training and, in particular, teacher job experience are highly significant predictors of the teacher's statute in both grades.

4.3. Matching results for the effect of the contract teacher statute

Table 4 presents the results of a comparison between the scores of classes with a contract teacher with the counterfactual outcome of the same classes if they had had a traditional teacher. Estimates are presented for both specifications by subject and grade. The estimation procedure imposes common support on the basis of the propensity score. 90% confidence intervals are estimated using bootstrapped standard errors (500 replications). The estimated bias is calculated from the bootstrapped distribution and indicates to what extent the normality assumption underlying the estimation of the confidence interval is actually adequate.

Table 4: Estimation of the effect of the contract teacher statute (ATT)[1]

	Grade 2						Grade 5					
Specification	N[2]	ATT[3]	Bias	Std. Err.[4]	[90% Conf.	Interval]	N[2]	ATT[3]	Bias	Std. Err.[4]	[90% Conf.	Interval]
1 Math+French	123 (121)	**-5.43**	0.69	2.80	-10.04	-0.81	117 (114)	-0.67	-0.01	1.56	-3.25	1.90
1 French	123 (120)	**-6.52**	0.96	2.88	-11.27	-1.77	117 (114)	-0.23	-0.07	1.63	-2.93	2.46
1 Math	123 (121)	-5.03	0.65	3.29	-10.46	0.39	117 (113)	-0.44	-0.62	1.79	-3.39	2.51
2 Math+French	97 (89)	0.83	-0.64	5.43	-8.11	9.77	69 (56)	3.09	-1.21	3.63	-2.90	9.08
2 French	97 (89)	1.93	-1.30	6.27	-8.41	12.27	69 (56)	4.40	-1.81	3.96	-2.12	10.94
2 Math	97 (90)	-0.80	-0.29	6.00	-12.80	7.00	69 (58)	2.18	-0.95	3.77	-4.04	8.40

[1] Figures computed using PSmatch2 (Leuwen and Sianesi 2003).
[2] Number of observations after imposing common support in brackets.
[3] Bold figures indicate significance at the 10% level.
[4] Calculated on the basis of 500 bootstrap replications.

Table 4 shows that in 5th grade, the control for inappropriate comparisons implied by the matching procedure considerably reduces the observed underperformance of contract teachers (as compared to the gross differences observed in Table 2). This difference is no longer significant, not even at the 10% level. In specification 2, the coefficient even becomes positive – although it remains insignificant.

As expected, matching produces much less effect in 2nd grade. The differences in scores observed earlier in Table 2 are reproduced and even reinforced here. These differences (to the advantage of the traditional teacher statute) are not only significant but also sizeable. The observed increase in performance of 6 percentage points corresponds to almost 40% of a standard deviation and to 15% of the mean score (average for Math and French across all 2nd grade classes).[9] A potential explanation for the strong differences between the results for grade 2 and grade 5 may be that out of the group of contract teachers, only the best are employed for classes in higher grades, which are often considered as more important by both parents and principals.

However, once job experience and teacher's education and training are introduced as additional controls (specification 2), even in grade 2 the difference between both types of teachers is no more significant. For French, as well as French and Math combined, the point estimates even switch signs. These results are relatively robust with respect to the inclusion of other controls.

[9] See Annex 3 for the general statistics on the distribution of scores.

Obviously, for both grades, specification 2 suffers from the limited number of observations in the restricted sample. Moreover, the estimated bias reveals that the bootstrap distribution is not well centered around the mean so that the estimated confidence interval may not be truly reliable. However, adjusting the confidence interval to the bias does not change our results.

Annex 4 provides an overview of the effect of matching on the balancing of characteristics between the groups of students taught by contract teachers and civil servants respectively. It confirms once again that with respect to the variables considered in specification 1, matching appears to be relevant primarily in grade 5, while in grade 2, characteristics are relatively well balanced right from the beginning. However, once teacher job experience and training are introduced as additional controls for the reduced sample (specification 2), results of the balancing statistics clearly reveal the relevance of matching for both grades.

All in all, our final results on the impact of the contract teacher statute do not confirm the concerns about a potential deterioration of education quality. The overall assessment of the contract teacher program in Niger is clearly positive for grade 5, since the positive effects on enrolment are not diminished by any significant negative effect on student performance, whichever model specification is used. In grade 2, at first glance, we seem to face a quantity-quality trade-off, but the limitation of the sample to younger teachers and the control for job experience and teacher training lead to different results according to which, once again, the teacher statute does not seem to play any significant role.

Nevertheless, even if one was to believe the results of the full sample without control for teacher job experience rather than the results of the reduced sample, the comparison of the losses due to reduced quality and the gains due to increased enrolment might still turn into a positive overall assessment. Assuming a fixed budget for primary education and taking into account that public contract teachers earn (less than) 50% of civil servant teacher salaries (Table 2), only about half the number of traditional teachers could have been employed. Ceteris paribus, given that public contract teachers represented 50% of all teachers in 2000 (Table 1), this implies that without the contract teacher program, the total number of teachers would have been only 75% of what it effectively was. At unchanged class size, this would have implied that, correspondingly, only 75% of today's students would have been enrolled. Now the potential reduction in achievement of 5.4 percentage points for all 2nd grade students can be compared to a situation where 75% of the students do not face these quality losses, but 25% are not enrolled at all. In this case, a positive assessment of the contract teacher program is obtained whenever the performance of unenrolled children is expected to lie at least one standard deviation below the performance of those enrolled.[10]

Finally, it must be considered that those who would not have been enrolled typically belong to the most disadvantaged (i.e. poorest) groups of the society. This implies that the contract teacher program has had an additional positive effect on distribution and on the benefit incidence of public expenditure on education for the poor.

5. Experiences from other countries and conclusions

All in all, this paper has shown that contract teacher programs have enabled Niger and other countries to considerably enhance enrolment. The introduction of these programs therefore represented an important step towards universal primary enrolment and completion as agreed

[10] This result is obtained by computing $0.75 \times 5.4 / 0.25 \times 16.3 = 0.99$, where 5.4 is the point estimate for Math and French combined using specification 1, and 16.3 is the standard deviation of test scores.

upon internationally in the Millennium Development Goals and the Education for All objectives. From this perspective, they also represent a major step towards the reduction of poverty. Nevertheless, a certain quantity-quality trade-off might exist and has to be taken seriously, as schooling will only enhance children's abilities to master their everyday lives if at least a certain minimum quality can be ensured. For the specific case of Niger, we find a significant deterioration of education quality measured in terms of student achievement only for grade 2, and even there, only as long as we do not restrict our sample to younger teachers so that we can control for job experience. Comparing the different specifications, it appears that the negative effect of contract teachers in the full sample is largely an artifact of their limited job experience. Moreover, in any case, the positive impact on enrolment seems to dominate over potential losses in quality.

This overall assessment is in line with the conclusions reached by PASEC (2005a) - despite the fact that the authors tend to find a significant negative effect in 5th rather than in 2nd grade. However, the simple (clustered) linear regression analysis used seems to lead to somewhat less robust results - as indicated by the coefficients and varying significance levels of various specifications presented in the annex of their study.

Similar results have also been found for Togo and Guinea. In Togo, the effect of contract teachers on student achievement was shown to be significantly negative for both grades, and in Guinea, a significantly negative impact was found for grade 2 (PASEC 2005a, 2004, 2003, and Vegas and de Laat 2003). At the same time, the PASEC studies suggest that this negative impact may be driven to some extent by the impact of reduced teacher training. In fact, if the sample in Togo is split between teachers with and without initial training, the contract teacher statute is no more significant in the former. While Guinean teachers all receive some training, its structure and duration was reformed from the first to the second cohort, and indeed, the second cohort does not have any significantly negative (but partially even a slight positive effect) on students' learning.

Interestingly, in some countries, contract teacher programs have also led to an obvious improvement of results. A positive impact of contract teachers, significant for grade 2, was found in Mali, and this result was rather robust to the use of different estimation methods and model specifications (PASEC 2005b, Bourdon 2005). Moreover, an early study on the five countries Burkina Faso, Cameroon, Côte d'Ivoire, Madagascar and Senegal also found a significant and positive effect (Michaelowa 2001b). As opposed to Niger, where virtually all contract teachers are employed by public authorities, in all these cases the majority of contract teachers were engaged by parents or local communities. It seems that these teachers are relatively highly motivated and also miss classes less often than their peers employed in the civil service (Michaelowa 2002).

This suggests the relevance of the impact of this new type of contracts on incentives. While a teacher engaged as a civil servant faces hardly any incentives to increase his or her effort, the situation is different for teachers engaged on a short-term contract which may or may not be prolonged. But more importantly, if teachers are directly engaged by parents or the local community, they will feel the recognition of their effort and they can also be very directly held responsible for their work. The evidence available so far suggests that this may easily outweigh lower pay and adverse working conditions.

As contract teacher programs have been shown to be highly relevant tools in progress towards Education for All, towards a strengthening of capabilities and hence, towards poverty reduction, it seems to be important to further examine under which conditions their positive

impact on enrolment can go hand in hand with an equally positive effect on education quality. While contract teacher programs appear to have been worthwhile policies in all countries, the quality-quantity trade-off apparent in some countries might still be considerably reduced. This requires further research based on a distinction between the different types of contract teachers.

References

Amelewonou, Kokou, Mathieu Brossard and Luc Gacougnolle (2004): La question enseignante dans la perspective de la scolarisation primaire universelle en 2015 dans les pays de la CEDEAO, CEMAC et PALOP, Dakar: Pôle de Dakar – Coopération française/UNESCO (http://www.poledakar.org/article.php?id_article=112, 21/03/05)

Bernard, Jean-Marc (1999). Les Enseignants du primaire dans Cinq Pays du Programme d'Analyse des Systèmes Educatifs de la CONFEMEN : Le Rôle du Maître dans le Processus d'Acquisition des Élèves, Report for the Working Group on the Teaching Profession, ADEA, French Section, Paris

Bernard, Jean-Marc, Beïfith Kouak Tiyab and Katia Vianou (2004): Profils enseignants et qualité de l'éducation primaire en Afrique subsaharienne francophone: Bilan et perspectives de dix années de recherche du PASEC, mimeo, PASEC/CONFEMEN

Black, Dan and Jeffrey Smith (2004): How Robust is the Evidence on the Effects of College Quality? Evidence from Matching, Journal of Econometrics Vol. 121, No. 1-2, pp. 99-124.

Bourdon, Jean (2005): Les maîtres contractuels sont-ils aussi efficaces que les maîtres titulaires? Essai d'appariement sur les données de l'enquête PASEC sur le Mali, mimeo, IREDU, University of Burgundy

Bruns, Barbara, Alain Mingat and Ramahatra Rakotomalala (2002): A Chance for Every Child, Washington: The World Bank

Faoura, Thiam (ed.) (2004): La réforme de la formation initiale des maîtres en Guinée (FIMG): étude-bilan de la mise en oeuvre, Paris: L'Harmattan (ADEA publication)

Heckman, James and Richard Robb (1985): Alternative Methods for Evaluating the Impact of Interventions, in: James Heckman and Burton Singer (eds): Longitudinal Analysis of Labour Market Data, Cambridge: Cambridge University Press, pp. 156-245

Lechner, Michael (1999): Earnings and Employment Effects of Continuous Off-the Job Training in East Germany after Unification, in: Journal of Business and Economic Statistics Vol. 17, No. 1, pp. 74-90

Leuven, Eric and Barbara Sianesi (2003): PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing (http://ideas.repec.org/boc/bocode/s432001.html., 1/4/05)

MEB_1A (2004): Statistiques de l'éducation 2003-2004, Niamey: Ministère de l'Éducation de base 1 et de l'alphabétisation, République du Niger

Michaelowa, Katharina (2001a): Die Bedeutung von Bildung in Niedrigeinkommensländern: Das Beispiel Afrika, in: Renate Schubert (ed.): Entwicklungsperspektiven von Niedrigeinkommensländern – zur Bedeutung von Wissen und Institutionen, Berlin: Duncker&Humblot, pp. 127-152

Michaelowa, Katharina (2001b): Primary Education Quality in Francophone Sub-Saharan Africa: Determinants of Learning Achievement and Efficiency Considerations, in: World Development, Vol. 29, No. 10, pp. 1699-1716

Michaelowa, Katharina (2002): Teacher Job Satisfaction, Student Achievement, and the Cost of Primary Education in Francophone Sub-Saharan Africa, HWWA Discussion Paper No. 188, Hamburg

MINEDAF (2002): Universal Primary Education: Goal for All, statistical document, Conference of the Ministers of Education of African Member States (MINEDAF VIII), Dar-Es-Salaam: UNESCO

Mingat, Alain and Bruno Suchaut (2000): Une analyse économique comparative des systèmes éducatifs africains, Paris: De Boeck

OECD (2002): Education at a Glance, Paris

PASEC (2003): Le Programme de formation initiale des maîtres et la double vacation en Guinée, Dakar: CONFEMEN

PASEC (2005a): Les enseignants contractuels et la qualité de l'enseignement de base I au Niger : Quel bilan?, Dakar : CONFEMEN, forthcoming

PASEC (2005b): Impact du statut enseignant sur les acquisitions dans le premier cycle l'enseignement fondamental public au Mali, Dakar : CONFEMEN, forthcoming

PASEC (2004): Recrutement et formation des enseignants du premier degré au Togo: Quelles priorités?, Dakar : CONFEMEN

Rivkin, Steven, Eric Hanushek and John Kain (2005): Teachers, Schools, and Academic Achievement, Econometrica Vol. 73, No. 2, pp. 417-458

Rosenbaum, Paul and Donald Rubin (1983): The Central Role of the Propensity Score in Observational Studies for Causal Effects, in : Biometrika Vol. 70, No. 1, pp. 41-55

Vegas, Emiliana and Joost de Laat (2003): Do Differences in Teacher Contracts Affect Student Performance? Evidence from Togo, Working Paper No. 26955, The World Bank

UIS (2005): Education : Statistical Tables (http://portal.unesco.org/uis/ev.php?URL_ID=5187&URL_DO=DO_TOPIC&URL_SECTION=201, 21/3/05)

UNESCO (2004): Education for All – The Quality Imperative, EFA Global Monitoring Report 2005, Paris: UNESCO

ANNEX

Annex 1: Relationship between final test scores and teacher job experience

Figure A1.1: Partial scatter plot of final test scores (class average) on teacher job experience, for classes taught by civil servant teachers, grade 2

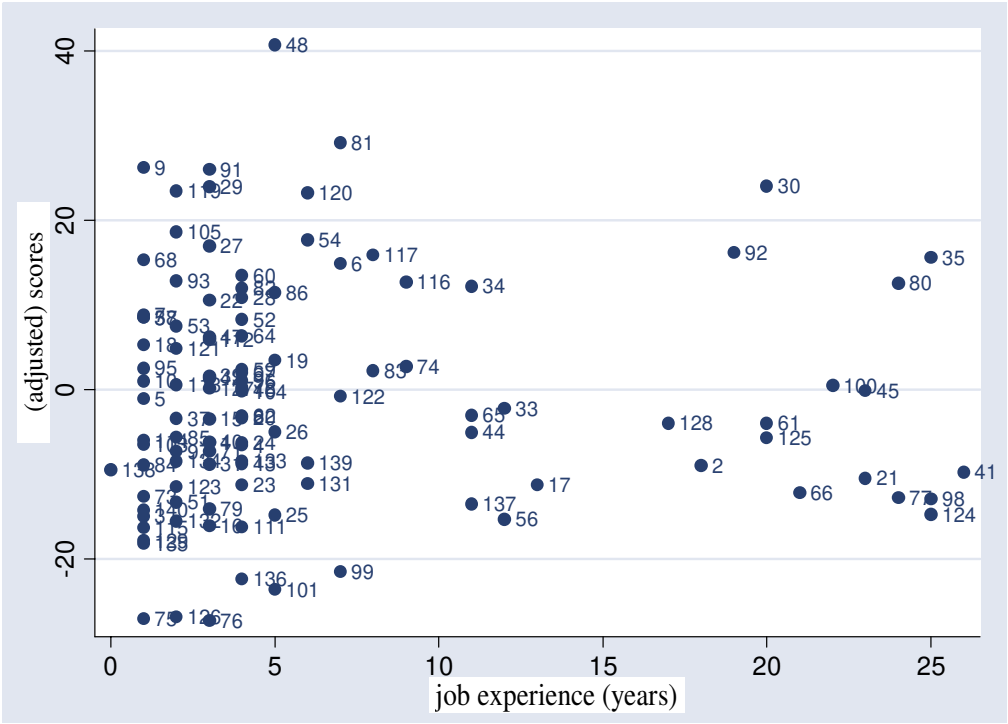

Figure A1.2: Partial scatter plot of final test scores (class average) on teacher job experience, for classes taught by civil servant teachers, grade 5

Notes: The partial scatter plots reflect the relationship between class level student achievement and job experience once all other variables are controlled for which satisfy the condition of a stepwise backward selection procedure with a significance level of 20% for removal from the model. Labels refer to school identification numbers.

Annex 2: Estimation of the propensity score, by subject taught

Table A2.1: French[1]

Probit estimates	Grade 2 Specification 1		Grade 2 Specification2		Grade 5 Specification 1		Grade 5 Specification2	
Dependent variable: Teacher is a contract teacher	Coef.	P>|z|	Coef.	P>|z|	Coef.	P>|z|	Coef.	P>|z|
Initial score, French	-0.01	0.36	-0.01	0.34	-0.02	0.29	-0.02	0.45
Index of parents' literacy (0-2)[2]	**-0.67**	**0.06**	**-1.31**	**0.01**	**0.90**	**0.05**		
Average number of textbooks per student (0-2)	**0.89**	**0.05**	**1.76**	**0.01**				
Share of Tamacheq speaking students					**5.77**	**0.02**	**9.44**	**0.07**
School is participating in a pilot project					**-0.71**	**0.03**	**-0.89**	**0.04**
Class size	0.04	0.19						
Class size squared	-0.00	0.13						
School is situated in a rural area					**0.86**	**0.03**		
Share of multi-grade classes in the region					**-6.77**	**0.05**	-6.73	0.18
Teacher's job experience (in years)			**-0.63**	**0.00**			**-0.33**	**0.01**
Index of teacher's educational attainment (0-6)			-0.25	0.17			-0.37	0.14
Duration of teacher training index (0-4)			**0.36**	**0.06**			**-0.54**	**0.04**
Constant	-0.24	0.77	**3.55**	**0.00**	-0.80	0.21	**4.37**	**0.01**
N	123		97		117		69	
Pseudo R^2	0.06		0.37		0.18		0.37	
LR, Prob > chi^2	0.09		0.00		0.00		0.00	

Table A2.2: Mathematics[1]

Probit estimates	Grade 2 Specification 1		Grade 2 Specification2		Grade 5 Specification 1		Grade 5 Specification2	
Dependent variable: Teacher is a contract teacher	Coef.	P>|z|	Coef.	P>|z|	Coef.	P>|z|	Coef.	P>|z|
Initial score, Mathematics	**-0.01**	**0.09**	**-0.01**	**0.10**	-0.02	0.24	-0.02	0.36
Index of parents' literacy (0-2)[2]	**-0.69**	**0.05**	**-1.34**	**0.01**	**0.97**	**0.04**		
Average number of textbooks per student (0-2)	**0.98**	**0.03**	**1.86**	**0.01**				
Share of Tamacheq speaking students					**5.68**	**0.01**	**9.87**	**0.06**
School is participating in a pilot project					**-0.64**	**0.05**	**-0.74**	0.11
Class size	0.04	0.19						
Class size squared	-0.00	0.11						
School is situated in a rural area					**0.82**	**0.03**		
Share of multi-grade classes in the region					**-6.28**	**0.07**	-6.78	0.17
Teacher's job experience (in years)			**-0.64**	**0.00**			**-0.32**	**0.01**
Index of teacher's educational attainment (0-6)			-0.30	0.12			-0.35	0.16
Duration of teacher training index (0-4)			**0.40**	**0.05**			**-0.53**	**0.05**
Constant	-0.01	0.99	**3.98**	**0.00**	-0.81	0.20	**4.38**	**0.00**
N	123		97		117		69	
Pseudo R^2	0.07		0.39		0.18		0.37	
LR, Prob > chi^2	0.04		0.00		0.00		0.00	

[1] Bold figures indicate significance at the 10% level.
[2] For grade 2 this variable does not exist. However, assuming that parents' literacy level does not differ much within a given school, the missing information was approximated by information available for grade 5 (see also footnote 7).

Annex 3: The distribution of test scores

Table A3.1: The distribution of class average test scores by subject and grade

	Grade 2		Grade 5	
	Mean	Std. deviation	Mean	Std. deviation
Math+French	41.1	16.3	29.9	9.4
French	40.5	17.9	28.5	10.1
Mathematics	39.7	16.4	31.2	9.9

Annex 4: Balancing characteristics of treated and controls before and after matching[1, 2] (Mathematics and French combined)

Table A4.1: Comparison of means, grade 2, specification 1

	Mean			%reduct	t-test	
Variable and sample	Treated	Control	%bias	|bias|	t	p>|t|
Initial score, average Math and French						
Unmatched	22.53	25.66	-19.2		-1.05	0.29
Matched	22.53	24.87	-14.3	25.2	-0.82	0.42
Average number of textbooks per student (0-2)						
Unmatched	0.33	0.31	7		0.39	0.70
Matched	0.33	0.35	-7.5	-6.7	-0.38	0.71
Index of parents' literacy (0-2)						
Unmatched	0.82	0.92	-26.9		-1.48	0.14
Matched	0.82	0.86	-10.1	62.5	-0.53	0.60
Class size						
Unmatched	43.43	44.83	-9.4		-0.52	0.61
Matched	43.43	42.79	4.3	54.2	0.27	0.79
Class size squared						
Unmatched	2050.10	2281.70	-16.7		-0.92	0.36
Matched	2050.10	1999.90	-3.6	78.3	0.24	0.81

Table A4.2: Comparison of means, grade 2, specification 2

	Mean			%reduct	t-test	
Variable and sample	Treated	Control	%bias	|bias|	t	p>|t|
Teacher's job experience (in years)						
Unmatched	2.64	4.47	-117		-5.61	0.00
Matched	2.64	2.70	-4.1	96.5	-0.2	0.84
Index of teacher's educational index (0-4)						
Unmatched	3.82	4.03	-23		-1.04	0.30
Matched	3.82	4.05	-25.5	-10.8	-1.23	0.22
Duration of teachers training index (0-4)						
Unmatched	1.67	1.65	2.5		0.11	0.91
Matched	1.67	1.46	21.2	-744.9	1.05	0.30
Initial score, average Math and French						
Unmatched	24.02	27.47	-20.9		-0.95	0.35
Matched	24.02	24.22	-1.2	94.3	-0.06	0.95
Average number of textbooks per student (0-2)						
Unmatched	0.34	0.34	0		0.00	1.00
Matched	0.34	0.38	-14.5	-86449.8	-0.66	0.51
Index of parents' literacy (0-2)						
Unmatched	0.81	0.90	-21.4		-0.98	0.33
Matched	0.81	0.73	21.9	-2.5	0.99	0.32

Table A4.3: Comparison of means, grade 5, specification 1

	Mean			%reduct	t-test	
Variable and sample	Treated	Control	%bias	|bias|	t	p>|t|
Initial score, average Math and French						
Unmatched	25.93	28.23	-25.3		-1.11	0.27
Matched	25.93	26.24	-3.4	86.6	-0.19	0.85
Index of parents' literacy (0-2)						
Unmatched	0.93	0.84	22.5		1.02	0.31
Matched	0.93	0.94	-3.3	85.2	-0.17	0.87
Share of Tamacheq speaking students						
Unmatched	0.03	0.02	21.5		1.00	0.32
Matched	0.03	0.03	6.4	70.3	0.31	0.76
School is participating in a pilot project						
Unmatched	0.20	0.43	-49.9		-2.09	0.04
Matched	0.20	0.26	-12.6	74.7	-0.72	0.47
School is located in a rural area						
Unmatched	0.40	0.30	20.1		0.91	0.37
Matched	0.40	0.36	7.9	60.5	0.42	0.68
Share of muli-grade schools						
Unmatched	0.07	0.09	-39.3		-1.56	0.12
Matched	0.07	0.07	4.3	89.0	0.26	0.80

Table A4.4: Comparison of means, grade 5, specification 2

	Mean			%reduct	t-test	
Variable and sample	Treated	Control	%bias	|bias|	t	p>|t|
Teacher's job experience (in years)						
Unmatched	3.47	4.34	-46.4		-1.62	0.11
Matched	3.47	4.00	-28.6	38.5	-1.06	0.29
Index of teacher's educational attainment (0-6)						
Unmatched	4.67	4.56	12.5		0.39	0.70
Matched	4.67	4.46	25	-99.7	0.93	0.36
Duration of teacher training index (0-4)						
Unmatched	1.27	1.95	-80.4		-2.39	0.02
Matched	1.27	1.42	-17.9	77.7	-0.84	0.40
Initial score, average Math and French						
Unmatched	27.69	27.06	6.9		0.23	0.82
Matched	27.69	26.65	11.4	-64.2	0.38	0.70
Share of Tamacheq speaking students						
Unmatched	0.02	0.01	28.6		1.05	0.30
Matched	0.02	0.02	-3.6	87.3	-0.11	0.91
School is participating in a pilot project						
Unmatched	0.33	0.41	-16.5		-0.54	0.59
Matched	0.33	0.28	11.2	31.8	0.44	0.66
Share of mulit-grade schools						
Unmatched	0.07	0.09	-34.2		-1.05	0.30
Matched	0.07	0.06	14.4	57.5	0.63	0.53

[1] Means for the matched groups are computed after imposing common support.
[2] Figures are computed using PSmatch2 (Leuwen and Sianesi 2003).